GI5202-CIP

TITLE OF THE INVENTION

TENDON-INDUCING COMPOSITIONS

RELATED APPLICATIONS

The present invention is a continuation-in-part of application serial number 08/164,103, filed on December 7, 1993.

FIELD OF THE INVENTION

The present invention relates to a novel family of purified proteins, and compositions containing such proteins, which compositions are useful for the induction of tendon/ligament-like tissue formation, wound healing and ligament and other tissue repair. These proteins may also be used in compositions for augmenting the activity of bone morphogenetic proteins.

BACKGROUND OF THE INVENTION

The search for the molecule or molecules responsible for formation of bone, cartilage, tendon and other tissues present in bone and other tissue extracts has led to the discovery of a novel set of molecules called the Bone Morphogenetic Proteins (BMPs). The structures of several proteins, designated BMP-1 through BMP-10, have previously been elucidated. The unique inductive activities of these proteins, along with their presence in bone, suggests that they are important regulators of bone repair processes, and may be involved in the normal maintenance of bone tissue. There is a need to identify additional proteins which play a role in forming other vital tissues. The present invention relates to the identification of a family of proteins, which have tendon/ligament-like tissue inducing activity, and which are useful in compositions for the induction of tendon/ligament-like tissue formation and repair.

SUMMARY OF THE INVENTION

In one embodiment, the present invention comprises DNA molecules encoding a tendon/ligament-like inducing protein

which the inventors have named V1-1. This novel protein is also called BMP-12. The present invention also includes DNA molecules encoding V1-1 related proteins.

V1-1 related proteins are a subset of the BMP/TGF-β/Vg-1 family of proteins, including V1-1 and VL-1, which are defined as tendon/ligament-like tissue inducing proteins encoded by DNA sequences which are cloned and identified, e.g., using PCR, using V1-1 specific primers, such as primers #6 and #7 described below, with reduced stringency conditions. It is preferred that the DNA sequences encoding V1-1 related proteins share at least about 80% homology at the amino acid level from amino acids with amino acids #3 to #103 of SEQ ID NO:1.

The DNA molecules preferably have a DNA sequence encoding the V1-1 protein, the sequence of which is provided in SEQ ID NO:1, or a V1-1 related protein as further described herein. Both the V1-1 protein and V1-1 related proteins are characterized by the ability to induce the formation of tendon/ligament-like tissue in the assay described in the examples.

The DNA molecules of the invention preferably comprise a DNA sequence, as described in SEQUENCE ID NO:1; more preferably nucleotides #571 to #882 or nucleotides #577 to #882 of SEQ ID NO:1; or DNA sequences which hybridize to the above under stringent hybridization conditions and encode a protein which exhibits the ability to form tendon/ligament-like tissue.

The DNA molecules of the invention also include DNA molecules comprising a DNA sequence encoding a V1-1 related protein with the amino acid sequence shown in SEQ ID NO:2, as well as naturally occurring allelic sequences and equivalent degenerative codon sequences of SEQ ID NO:2. Preferably, the DNA sequence of the present invention encodes amino acids #1 to # 104 or #3 to #103 of SEQ ID NO:2. The DNA sequence may comprise, in a 5′ to 3′ direction, nucleotides encoding a propeptide, and nucleotides encoding for amino acids #1 to #104 or #3 to #103 of SEQ ID NO:2. The propeptide useful in the

above embodiment is preferably selected from the group consisting of native V1-1 propeptide and a BMP protein propeptide. The invention further comprises DNA sequences which hybridize to the above DNA sequences under stringent hybridization conditions and encode a V1-1 related protein which exhibits the ability to induce formation of tendon/ligament-like tissue.

In other embodiments, the present invention comprises host cells and vectors which comprise a DNA molecule encoding the V1-1 protein, or a V1-1 related protein. The host cells and vectors may further comprise the coding sequence in operative association with an expression control sequence therefor.

In another embodiment, the present invention comprises a method for producing a purified V1-1 related protein, said method comprising the steps of culturing a host cell transformed with the above DNA molecule or vector comprising a nucleotide sequence encoding a V1-1 related protein; and (b) recovering and purifying said V1-1 related protein from the culture medium. In a preferred embodiment, the method comprises (a) culturing a cell transformed with a DNA molecule comprising the nucleotide sequence from nucleotide #571 or #577 to #879 or #882 as shown in SEQ ID NO:1; and

(b) recovering and purifying from said culture medium a protein comprising the amino acid sequence from amino acid #1 or #3 to amino acid #103 or #104 as shown in SEQ ID NO:2. The present invention also includes a purified protein produced by the above methods.

The present invention further comprises purified V1-1 related protein characterized by the ability to induce the formation of tendon/ligament-like tissue. The V1-1 related polypeptides preferably comprise an amino acid sequence as shown in SEQ ID NO:2. The polypeptide more preferably comprise amino acids #1 to #104 as set forth in SEQ ID NO:2; or amino acids #3 to #103 as set forth in SEQ ID NO:2. In a preferred embodiment, the purified polypeptide may be in the form of a

dimer comprised of two subunits, each with the amino acid sequence of SEQ ID NO:2.

In another embodiment, the present invention comprises compositions comprising an effective amount of the above-described V1-1 related proteins. In the compositions, the protein may be admixed with a pharmaceutically acceptable vehicle.

The invention also includes methods for tendon/ligament-like tissue healing and tissue repair, for treating tendinitis, or other tendon or ligament defects, and for inducing tendon/ligament-like tissue formation in a patient in need of same, comprising administering to said patient an effective amount of the above composition.

Other embodiments include chimeric DNA molecules comprising a DNA sequence encoding a propeptide from a member of the TGF-β superfamily of proteins linked in correct reading frame to a DNA sequence encoding a V1-1 related polypeptide. One suitable propeptide is the propeptide from BMP-2. The invention also includes heterodimeric protein molecules comprising one monomer having the amino acid sequence shown in SEQ ID NO:2, and one monomer having the amino acid sequence of another protein of the TGF-β subfamily.

Finally, the present invention comprises methods for inducing tendon/ligament-like tissue formation in a patient in need of same comprising administering to said patient an effective amount of a composition comprising a protein which exhibits the ability to induce formation of tendon/ligament-like tissue, said protein having an amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4. The amino acid sequences are more preferably one of the following: (a) amino acids #1 or #3 to #103 or #104 of SEQ ID NO:2; (b) amino acids #1 or #19 to #119 or #120 of SEQ ID NO:4; (c) mutants and/or variants of (a) or (b) which exhibit the ability to form tendon and/or ligament. In other embodiments of the above method, the protein is encoded by a DNA sequence of SEQ ID NO:1 or SEQ ID

NO:3, more preferably one of the following: (a) nucleotides #571 or #577 to #879 or #882 of SEQ ID NO:1; (b) nucleotides #845 or #899 to #1201 or #1204 of SEQ ID NO:3; and (c) sequences which hybridize to (a) or (b) under stringent hybridization conditions and encode a protein which exhibits the ability to form tendon/ligament-like tissue.

Description of the Sequences

SEQ ID NO:1 is the nucleotide sequence encoding the human V1-1.

SEQ ID NO:2 is the amino acid sequence comprising the mature human V1-1 polypeptide.

SEQ ID NO:3 is the nucleotide sequence encoding the protein MP52.

SEQ ID NO:4 is the amino acid sequence comprising the mature MP52 polypeptide.

SEQ ID NO:5 is the nucleotide sequence of a specifically amplified portion of the human V1-1 encoding sequence.

SEQ ID NO:6 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:5.

SEQ ID NO:7 is the nucleotide sequence of a specifically amplified portion of the human VL-1 encoding sequence.

SEQ ID NO:8 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO:7.

SEQ ID NO:9 is the nucleotide sequence of the plasmid pALV1-781, used for expression of V1-1 in *E. coli*.

SEQ ID NO:10 is the nucleotide sequence of a fragment of the murine clone, mV1.

SEQ ID NO:11 is the amino acid sequence of a fragment of the murine protein encoded by mV1.

SEQ ID NO:12 is the nucleotide sequence of a fragment of the murine clone, mV2.

SEQ ID NO:13 is the amino acid sequence of a fragment of the murine protein encoded by mV2.

SEQ ID NO:14 is the nucleotide sequence of a fragment of the murine clone, mV9.

SEQ ID NO:15 is the amino acid sequence of a fragment of the murine protein encoded by mV9.

SEQ ID NO:16 is the amino acid sequence of a BMP/TGF-β/Vg-1 protein consensus sequence. The first Xaa represents either Gln or Asn; the second Xaa represents either Val or Ile.

SEQ ID NO:17 is the nucleotide sequence of oligonucleotide #1.

SEQ ID NO:18 is the amino acid sequence of a BMP/TGF-β/Vg-1 protein consensus sequence. The Xaa represents either Val or Leu.

SEQ ID NO:19 is the nucleotide sequence of oligonucleotide #2.

SEQ ID NO:20 is the nucleotide sequence of oligonucleotide #3.

SEQ ID NO:21 is the nucleotide sequence of oligonucleotide #4.

SEQ ID NO:22 is the nucleotide sequence of oligonucleotide #5

SEQ ID NO:23 is the nucleotide sequence of oligonucleotide #6.

SEQ ID NO:24 is the nucleotide sequence of oligonucleotide #7.

Brief Description of the Figures

Figure 1 is a comparison of the human V1-1 and human MP52 sequences.

Detailed Description of the Invention

The DNA sequences of the present invention are useful for producing proteins which induce the formation of tendon/ligament-like tissue, as described further below. The DNA sequences of the present invention are further useful for isolating and cloning further DNA sequences encoding V1-1 related proteins with similar activity. These V1-1 related proteins may be homologues from other species, or may be related proteins within the same species.

Still, a further aspect of the invention are DNA sequences

coding for expression of a tendon/ligament-like tissue inducing protein. Such sequences include the sequence of nucleotides in a 5′ to 3′ direction illustrated in SEQ ID NO:1, DNA sequences which, but for the degeneracy of the genetic code, are identical to the DNA sequence SEQ ID NO:1, and encode the protein of SEQ ID NO:2. Further included in the present invention are DNA sequences which hybridize under stringent conditions with the DNA sequence of SEQ ID NO:1 and encode a protein having the ability to induce the formation of tendon or ligament. Preferred DNA sequences include those which hybridize under stringent conditions as described in Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), pages 387 to 389. Finally, allelic or other variations of the sequences of SEQ ID NO:1, whether such nucleotide changes result in changes in the peptide sequence or not, but where the peptide sequence still has tendon/ligament-like tissue inducing activity, are also included in the present invention.

The human V1-1 DNA sequence (SEQ ID NO:1) and amino acid sequence (SEQ ID NO:2) are set forth in the Sequence Listings. Another protein that is useful for the compositions and methods of the present invention is VL-1. VL-1 is a V1-1 related protein which was cloned using sequences from V1-1. A partial DNA sequence of VL-1 (SEQ ID NO:7) and the encoded amino acid sequence (SEQ ID NO:8) are set forth in the Sequence Listings.

The coding sequence of the mature human V1-1 protein appears to begin at nucleotide #571 and continues through nucleotide #882 of SEQ ID NO:1. The first cysteine in the seven cysteine structure characteristic of TGF-β proteins begins at nucleotide #577. The last cysteine ends at #879. The MP52 protein appears to begin at nucleotide #845 of SEQ ID NO:3 and continues through nucleotide #1204 of SEQ ID NO:3. The first cysteine of the seven cysteine structure characteristic of TGF-β proteins begins at nucleotide #899. The last cysteine ends at #1201.

As described earlier, V1-1 related proteins are a subset of the BMP/TGF-β/Vg-1 family of proteins, including V1-1 and VL-1, which can be defined as tendon/ligament-like tissue inducing proteins encoded by DNA sequences which can be cloned and identified, e.g., using PCR, using V1-1 specific primers, such as primers #6 and #7 described below, with reduced stringency conditions. It is preferred that DNA sequences of the present invention share at least about 80% homology at the amino acid level from amino acids with the DNA encoding amino acids #3 to #103 of SEQ ID NO:1. For the purposes of the present invention, the term V1-1 related proteins does not include the human MP52 protein. Using the sequence information of SEQ ID NO:1 and SEQ ID NO:3, and the comparison provided in Figure 1, it is within the skill of the art to design primers to the V1-1 sequence which will allow for the cloning of genes encoding V1-1 related proteins.

In order to produce the purified tendon/ligament-like tissue inducing proteins useful for the present invention, a method is employed comprising culturing a host cell transformed with a DNA sequence comprising a suitable coding sequence, particularly the DNA coding sequence from nucleotide #571 or # 577 to #879 or #882 of SEQ ID NO:1; and recovering and purifying from the culture medium a protein which contains the amino acid sequence or a substantially homologous sequence as represented by amino acids #1 or #3 to #103 or #104 of SEQ ID NO:2.

The human MP52 DNA is described in WO93/16099, the disclosure of which is incorporated herein by reference. However, this document does not disclose the ability of the protein to form tendon/ligament-like tissue, or its use in compositions for induction of tendon/ligament-like tissue. Human MP52 was originally isolated using RNA from human embryo tissue. The human MP52 nucleotide sequence (SEQ ID NO:3) and the encoded amino acid sequences (SEQ ID NO:4) are set forth in the Sequence Listings herein. The coding sequence of the

mature human MP52 protein is expected to begin at nucleotide #845 and continues through nucleotide #1204 of SEQ ID NO:3. Purified human MP52 proteins of the present invention may be produced by culturing a host cell transformed with a DNA sequence comprising the DNA coding sequence of SEQ ID NO:3 from nucleotide #845 to #1204, and recovering and purifying from the culture medium a protein which contains the amino acid sequence or a substantially homologous sequence as represented by amino acids #1 to #120 of SEQ ID NO:4. It is also expected that the amino acid sequence from amino acids #17 or #19 to #119 or #120 of SEQ ID NO:4 will retain activity. Thus, the DNA sequence from nucleotides #845, #893 or #899 to #1201 or #1204 are expected to encode active proteins.

For expression of the protein in mammalian host cells, the host cell is transformed with a coding sequence encoding a propeptide suitable for the secretion of proteins by the host cell is linked in proper reading frame to the coding sequence for the mature protein. For example, see United States Patent 5,168,050, the disclosure of which is hereby incorporated by reference, in which a DNA encoding a precursor portion of a mammalian protein other than BMP-2 is fused to the DNA encoding a mature BMP-2 protein. Thus, the present invention includes chimeric DNA molecules comprising a DNA sequence encoding a propeptide from a member of the TGF-β superfamily of proteins, is linked in correct reading frame to a DNA sequence encoding a tendon/ligament-like tissue inducing polypeptide. The term "chimeric" is used to signify that the propeptide originates from a different polypeptide than the encoded mature polypeptide. Of course, the host cell may be transformed with a DNA sequence coding sequence encoding the native propeptide linked in correct reading frame to a coding sequence encoding the mature protein shown in SEQ ID NO:2 or SEQ ID NO:4. The full sequence of the native propeptide may be determined through methods known in the art using the sequences disclosed in SEQ ID NO:1 or SEQ ID NO:3 to design a suitable probe for

identifying and isolating the entire clone.

The present invention also encompasses the novel DNA sequences, free of association with DNA sequences encoding other proteinaceous materials, and coding for expression of tendon/ligament-like tissue inducing proteins. These DNA sequences include those depicted in SEQ ID NO:1 in a 5′ to 3′ direction and those sequences which hybridize thereto under stringent hybridization conditions [for example, 0.1X SSC, 0.1% SDS at 65°C; see, T. Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), pages 387 to 389] and encode a protein having tendon/ligament-like tissue inducing activity.

Similarly, DNA sequences which code for proteins coded for by the sequences of SEQ ID NO:1, or proteins which comprise the amino acid sequence of SEQ ID NO:2, but which differ in codon sequence due to the degeneracies of the genetic code or allelic variations (naturally-occurring base changes in the species population which may or may not result in an amino acid change) also encode the tendon/ligament-like tissue inducing proteins described herein. Variations in the DNA sequences of SEQ ID NO:1 which are caused by point mutations or by induced modifications (including insertion, deletion, and substitution) to enhance the activity, half-life or production of the polypeptides encoded are also encompassed in the invention.

Another aspect of the present invention provides a novel method for producing tendon/ligament-like tissue inducing proteins. The method of the present invention involves culturing a suitable cell line, which has been transformed with a DNA sequence encoding a protein of the invention, under the control of known regulatory sequences. The transformed host cells are cultured and the proteins recovered and purified from the culture medium. The purified proteins are substantially free from other proteins with which they are co-produced as well as from other contaminants.

Suitable cells or cell lines may be mammalian cells, such

as Chinese hamster ovary cells (CHO). As described above, expression of protein in mammalian cells requires an appropriate propeptide to assure secretion of the protein. The selection of suitable mammalian host cells and methods for transformation, culture, amplification, screening, product production and purification are known in the art. See, e.g., Gething and Sambrook, Nature, 293:620-625 (1981), or alternatively, Kaufman et al, Mol. Cell. Biol., 5(7):1750-1759 (1985) or Howley et al, U.S. Patent 4,419,446. Another suitable mammalian cell line, which is described in the accompanying examples, is the monkey COS-1 cell line. The mammalian cell CV-1 may also be suitable.

Bacterial cells may also be suitable hosts. For example, the various strains of E. coli (e.g., HB101, MC1061) are well-known as host cells in the field of biotechnology. Various strains of B. subtilis, Pseudomonas, other bacilli and the like may also be employed in this method. For expression of the protein in bacterial cells, DNA encoding a propeptide is not necessary.

Bacterial expression of mammalian proteins, including members of the TGF-β family is known to produce the proteins in a non-glycosylated form, and in the form of insoluble pellets, known as inclusion bodies. Techniques have been described in the art for solubilizing these inclusion bodies, denaturing the protein using a chaotropic agent, and refolding the protein sufficiently correctly to allow for their production in a soluble form. For example, see EP 0433225, the disclosure of which is hereby incorporated by reference.

Alternatively, methods have been devised which circumvent inclusion body formation, such as expression of gene fusion proteins, wherein the desired protein is expressed as a fusion protein with a fusion partner. The fusion protein is later subjected to cleavage to produce the desired protein. One example of such a gene fusion expression system for E. coli is based on use of the E. coli thioredoxin gene as a fusion

partner, LaVallie et al., Bio/Technology, 11:187-193 (1993), the disclosure of which is hereby incorporated by reference.

Many strains of yeast cells known to those skilled in the art may also be available as host cells for expression of the polypeptides of the present invention. Additionally, where desired, insect cells may be utilized as host cells in the method of the present invention. See, e.g. Miller et al, Genetic Engineering, 8:277-298 (Plenum Press 1986) and references cited therein.

Another aspect of the present invention provides vectors for use in the method of expression of these tendon/ligament-like tissue inducing proteins. Preferably the vectors contain the full novel DNA sequences described above which encode the novel factors of the invention. Additionally, the vectors contain appropriate expression control sequences permitting expression of the protein sequences. Alternatively, vectors incorporating modified sequences as described above are also embodiments of the present invention. Additionally, the sequence of SEQ ID NO:1 or SEQ ID NO:3 could be manipulated to express a mature protein by deleting propeptide sequences and replacing them with sequences encoding the complete propeptides of BMP proteins or members of the TGF-β superfamily. Thus, the present invention includes chimeric DNA molecules encoding a propeptide from a member of the TGF-β superfamily linked in correct reading frame to a DNA sequence encoding a protein having the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:4.

The vectors may be employed in the method of transforming cell lines and contain selected regulatory sequences in operative association with the DNA coding sequences of the invention which are capable of directing the replication and expression thereof in selected host cells. Regulatory sequences for such vectors are known to those skilled in the art and may be selected depending upon the host cells. Such selection is routine and does not form part of the present invention.

A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells or induce differentiation of progenitors of tendon- or ligament-forming cells. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects.

The V1-1 related proteins may be recovered from the culture medium and purified by isolating them from other proteinaceous materials from which they are co-produced and from other contaminants present. The proteins of the present invention are capable of inducing the formation of tendon/ligament-like tissue. These proteins may be further characterized by the ability to demonstrate tendon/ligament-like tissue formation activity in the rat ectopic implant assay described below. It is contemplated that these proteins may have ability to induce the formation of other types of tissue, such as ligaments, as well.

The tendon/ligament-like tissue inducing proteins provided herein also include factors encoded by the sequences similar to

those of SEQ ID NO:1, but into which modifications are naturally provided (e.g. allelic variations in the nucleotide sequence which may result in amino acid changes in the polypeptide) or deliberately engineered. For example, synthetic polypeptides may wholly or partially duplicate continuous sequences of the amino acid residues of SEQ ID NO:2. These sequences, by virtue of sharing primary, secondary, or tertiary structural and conformational characteristics with tendon/ligament-like tissue growth factor polypeptides of SEQ ID NO:2 may possess tendon/ligament-like or other tissue growth factor biological properties in common therewith. Thus, they may be employed as biologically active substitutes for naturally-occurring tendon/ligament-like tissue inducing polypeptides in therapeutic compositions and processes.

Other specific mutations of the sequences of tendon/ligament-like tissue inducing proteins described herein involve modifications of glycosylation sites. These modifications may involve O-linked or N-linked glycosylation sites. For instance, the absence of glycosylation or only partial glycosylation results from amino acid substitution or deletion at asparagine-linked glycosylation recognition sites. The asparagine-linked glycosylation recognition sites comprise tripeptide sequences which are specifically recognized by appropriate cellular glycosylation enzymes. These tripeptide sequences may be asparagine-X-threonine, asparagine-X-serine or asparagine-X-cysteine, where X is usually any amino acid except proline. A variety of amino acid substitutions or deletions at one or both of the first or third amino acid positions of a glycosylation recognition site (and/or amino acid deletion at the second position) results in non-glycosylation at the modified tripeptide sequence. Additionally, bacterial expression of protein will also result in production of a non-glycosylated protein, even if the glycosylation sites are left unmodified.

The compositions of the present invention comprise a

purified V1-1 related protein which may be produced by culturing a cell transformed with the DNA sequence of SEQ ID NO:1 and recovering and purifying protein having the amino acid sequence of SEQ ID NO:2 from the culture medium. The purified expressed protein is substantially free from other proteinaceous materials with which it is co-produced, as well as from other contaminants. The recovered purified protein is contemplated to exhibit tendon/ligament-like tissue formation activity, and other tissue growth activity, such as ligament regeneration. The proteins of the invention may be further characterized by the ability to demonstrate tendon/ligament-like tissue formation activity in the rat assay described below.

The compositions for inducing tendon/ligament-like tissue formation of the present invention may comprise an effective amount of a tendon/ligament-like tissue inducing protein, wherein said protein comprises the amino acid sequence of SEQ ID NO:2, preferably amino acids #1 or #3 to #103 or #104 of SEQ ID NO:2; as well as mutants and/or variants of SEQ ID NO:2 which exhibit the ability to form tendon and/or ligament

Compositions of the present invention may further comprise additional proteins, such as additional members of the TGF-β superfamily of proteins, such as activins. Another aspect of the invention provides pharmaceutical compositions containing a therapeutically effective amount of a tendon/ligament-inducing protein, such as V1-1 or VL-1, in a pharmaceutically acceptable vehicle or carrier. These compositions may be used to induce the formation of tendon/ligament-like tissue or other tissue. It is contemplated that such compositions may also be used for tendon and ligament repair, wound healing and other tissue repair, such as skin repair. Compositions of the invention may further include at least one other therapeutically useful agent, such as the BMP proteins BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6 and BMP-7, disclosed for instance in United States Patents 5,108,922; 5,013,649;

5,116,738; 5,106,748; 5,187,076; and 5,141,905; BMP-8, disclosed in PCT publication WO91/18098; BMP-9, disclosed in PCT publication WO93/00432; and BMP-10, disclosed in co-pending patent application, serial number 08/061,695, filed on May 12, 1993. The disclosure of the above documents are hereby incorporated by reference herein.

The compositions of the invention may comprise, in addition to a tendon/ligament-inducing protein such as V1-1 or VL-1, other therapeutically useful agents including MP52, epidermal growth factor (EGF), fibroblast growth factor (FGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and fibroblast growth factor-4 (FGF-4), parathyroid hormone (PTH), leukemia inhibitory factor (LIF/HILDA/DIA), insulin-like growth factors (IGF-I and IGF-II). Portions of these agents may also be used in compositions of the present invention.

It is contemplated that the compositions of the invention may also be used in wound healing, such as skin healing and related tissue repair. The types of wounds include, but are not limited to burns, incisions and ulcers. (See, e.g. PCT Publication WO84/01106 for discussion of wound healing and related tissue repair).

It is expected that the proteins of the invention may act in concert with or perhaps synergistically with other related proteins and growth factors. Further therapeutic methods and compositions of the invention therefore comprise a therapeutic amount of at least one protein of the invention with a therapeutic amount of at least one of the BMP proteins described above. Such compositions may comprise separate molecules of the BMP proteins or heteromolecules comprised of different BMP moieties. For example, a method and composition of the invention may comprise a disulfide linked dimer comprising a V1-1 related protein subunit and a subunit from one of the "BMP" proteins described above. Thus, the present invention includes compositions comprising a purified V1-1

related polypeptide which is a heterodimer wherein one subunit comprises the amino acid sequence from amino acid #1 to amino acid #104 of SEQ ID NO:2, and one subunit comprises an amino acid sequence for a bone morphogenetic protein selected from the group consisting of BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, BMP-9 and BMP-10. A further embodiment may comprise a heterodimer of tendon/ligament-like tissue inducing moieties such as V1-1, VL-1 or MP52. For example the heterodimer may comprise one subunit comprising an amino acid sequence from #1 to # 104 of SEQ ID NO:2 and the one subunit may comprise an amino acid sequence from #1 to #120 of SEQ ID NO:4. Further, compositions of the present invention may be combined with other agents beneficial to the treatment of the defect, wound, or tissue in question.

The preparation and formulation of such physiologically acceptable protein compositions, having due regard to pH, isotonicity, stability and the like, is within the skill of the art. The therapeutic compositions are also presently valuable for veterinary applications due to the lack of species specificity in TGF-β proteins. Particularly domestic animals and thoroughbred horses in addition to humans are desired patients for such treatment with the compositions of the present invention.

The therapeutic method includes administering the composition topically, systemically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than the proteins which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the

methods of the invention.

The compositions may also include an appropriate matrix and/or sequestering agent as a carrier. For instance, the matrix may support the composition or provide a surface for tendon/ligament-like tissue formation and/or other tissue formation. The matrix may provide slow release of the protein and/or the appropriate environment for presentation thereof. The sequestering agent may be a substance which aids in ease of administration through injection or other means, or may slow the migration of protein from the site of application.

The choice of a carrier material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined. Potential carriers, such as sequestering agents, which may be suitable for the compositions of the present invention may be biodegradable, and may include alkylcellulosic materials.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate

handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

Additional optional components useful in the practice of the subject application include, e.g. cryogenic protectors such as mannitol, sucrose, lactose, glucose, or glycine (to protect the protein from degradation during lyophilization), antimicrobial preservatives such as methyl and propyl parabens and benzyl alcohol; antioxidants such as EDTA, citrate and BHT (butylated hydroxytoluene); and surfactants such as poly(sorbates) and poly(oxyethylenes); etc.

As described above, the compositions of the invention may be employed in methods for treating a number of tendon defects, such as the regeneration of tendon/ligament-like tissue in areas of tendon or ligament damage, to assist in repair of tears of tendon tissue, ligaments, and various other types of tissue defects or wounds. These methods, according to the invention, entail administering to a patient needing such tendon/ligament-like tissue or other tissue repair, a composition comprising an effective amount of a tendon/ligament-like tissue inducing protein, such as described in SEQ ID NO:2 and SEQ ID NO:4. These methods may also entail the administration of a tendon/ligament-like tissue inducing protein in conjunction with at least one of the BMP proteins described above. In another embodiment, the methods may entail administration of a heterodimeric protein in which one of the monomers is a tendon/ligament-like tissue inducing polypeptide, such as V1-1, VL-1 or MP52, and the second monomer is a member of the TGF-β superfamily of growth factors. In addition, these methods may also include the administration of a tendon/ligament-like tissue inducing protein with other growth factors including EGF, FGF, TGF-α, TGF-β, and IGF.

Thus, a further aspect of the invention is a therapeutic method and composition for repairing tendon/ligament-like

tissue, for repairing tendon or ligament as well as treating tendinitis and other conditions related to tendon or ligament defects. Such compositions comprise a therapeutically effective amount of one or more tendon/ligament-like tissue inducing proteins, such as V1-1, a V1-1 related protein, or MP52, in admixture with a pharmaceutically acceptable vehicle, carrier or matrix.

The dosage regimen will be determined by the attending physician considering various factors which modify the action of the composition, e.g., amount of tendon or ligament tissue desired to be formed, the site of tendon or ligament damage, the condition of the damaged tendon or ligament, the size of a wound, type of damaged tissue, the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and the types of additional proteins in the composition. The addition of other known growth factors, such as IGF-I (insulin like growth factor I), to the final composition, may also affect the dosage.

Progress can be monitored by periodic assessment of tendon/ligament-like tissue formation, or tendon or ligament growth and/or repair. The progress can be monitored by methods known in the art, for example, X-rays, arthroscopy, histomorphometric determinations and tetracycline labeling.

The following examples illustrate practice of the present invention in recovering and characterizing human tendon/ligament-like tissue inducing protein and employing them to recover the other tendon/ligament-like tissue inducing proteins, obtaining the human proteins, expressing the proteins via recombinant techniques, and demonstration of the ability of the compositions of the present invention to form tendon/ligament-like tissue in an <u>in vivo</u> model. Although the examples demonstrate the invention with respect to V1-1, with minor modifications within the skill of the art, the same results are believed to be attainable with MP52 and VL-1.

EXAMPLE 1

Isolation of DNA

DNA sequences encoding V1-1 and V1-1 related proteins may be isolated by various techniques known to those skilled in the art. As described below, oligonucleotide primers may be designed on the basis of amino acid sequences present in other BMP proteins, Vg-1 related proteins and other proteins of the TGF-β superfamily. Regions containing amino acid sequences which are highly conserved within the BMP family of proteins and within other members of the TGF-β superfamily of proteins can be identified and consensus amino acid sequences of these highly conserved regions can be constructed based on the similarity of the corresponding regions of individual BMP/TGF-β/Vg-1 proteins. An example of such a consensus amino acid sequence is indicated below.

Consensus amino acid sequence (1):

Trp-Gln/Asn-Asp-Trp-Ile-Val/Ile-Ala (SEQ ID NO:16)

Where X/Y indicates that either amino acid residue may appear at that position.

The following oligonucleotide is designed on the basis of the above identified consensus amino acid sequence (1):

#1: CGGATCCTGGVANGAYTGGATHRTNGC (SEQ ID NO:17)

This oligonucleotide sequence is synthesized on an automated DNA synthesizer. The standard nucleotide symbols in the above identified oligonucleotide primer are as follows: A,adenosine; C,cytosine; G,guanine; T,thymine; N,adenosine or cytosine or guanine or thymine; R,adenosine or cytosine; Y,cytosine or thymine; H,adenosine or cytosine or thymine; V,adenosine or cytosine or guanine; D,adenosine or guanine or thymine.

The first seven nucleotides of oligonucleotide #1 (underlined) contain the recognition sequence for the restriction endonuclease BamHI in order to facilitate the manipulation of a specifically amplified DNA sequence encoding the V1-1 protein and are thus not derived from the consensus

amino acid sequence (1) presented above.

A second consensus amino acid sequence is derived from another highly conserved region of BMP/TGF-β/Vg-1 proteins as described below:

His-Ala-Ile-Val/Leu-Gln-Thr (SEQ ID NO:18)

The following oligonucleotide is designed on the basis of the above identified consensus amino acid sequence (2):

#2: TTTCTAGAARNGTYTGNACDATNGCRTG (SEQ ID NO:19)

This oligonucleotide sequence is synthesized on an automated DNA synthesizer. The same nucleotide symbols are used as described above.

The first seven nucleotides of oligonucleotide #1 (underlined) contain the recognition sequence for the restriction endonuclease XbaI in order to facilitate the manipulation of a specifically amplified DNA sequence encoding the V1-1 protein and are thus not derived from the consensus amino acid sequence (2) presented above.

It is contemplated that the V1-1 protein of the invention and other BMP/TGF-β/Vg-1 related proteins may contain amino acid sequences similar to the consensus amino acid sequences described above and that the location of those sequences within a V1-1 protein or other novel related proteins would correspond to the relative locations in the proteins from which they were derived. It is further contemplated that this positional information derived from the structure of other BMP/TGF-β/Vg-1 proteins and the oligonucleotide sequences #1 and #2 which have been derived from consensus amino acid sequences (1) and (2),respectively, could be utilized to specifically amplify DNA sequences encoding the corresponding amino acids of a V1-1 protein or other BMP/TGF-β/Vg-1 related proteins.

Based on the knowledge of the gene structures of BMP/TGF-β/Vg-1 proteins it is further contemplated that human genomic DNA can be used as a template to perform specific amplification reactions which would result in the identification of V1-1 BMP/TGF-β/Vg-1 (V1-1 related protein) encoding sequences. Such

specific amplification reactions of a human genomic DNA template could be initiated with the use of oligonucleotide primers #1 and #2 described earlier. Oligonucleotides #1 and #2 identified above are utilized as primers to allow the specific amplification of a specific nucleotide sequence from human genomic DNA. The amplification reaction is performed as follows:

Human genomic DNA (source: peripheral blood lymphocytes), provided by Ken Jacobs of Genetics Institute, is sheared by repeated passage through a 25 gauge needle, denatured at 100°C for 5 minutes and then chilled on ice before adding to a reaction mixture containing 200 μM each deoxynucleotide triphosphates (dATP, dGTP, dCTP and dTTP), 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 1.25 units Taq DNA polymerase, 100 pM oligonucleotide #1 and 100 pM oligonucleotide #2. This reaction mixture is incubated at 94°C for two minutes and then subjected to thermal cycling in the following manner: 1 minute at 94°C, 1 minute at 40°C, 1 minute at 72°C for three cycles; then 1 minute at 94°C, 1 minute at 55°C, 1 minute at 72°C for thirty-seven cycles, followed by a 10 minute incubation at 72°C.

The DNA which is specifically amplified by this reaction is ethanol precipitated, digested with the restriction endonucleases BamHI and XbaI and subjected to agarose gel electrophoresis. A region of the gel, corresponding to the predicted size of the V1-1 or other BMP/TGF-β/Vg-1 encoding DNA fragment, is excised and the specifically amplified DNA fragments contained therein are electroeluted and subcloned into the plasmid vector pGEM-3 between the XbaI and BamHI sites of the polylinker. DNA sequence analysis of one of the resulting V1-1 related subclones indicates the specifically amplified DNA sequence product contained therein encodes a portion of the V1-1 protein of the invention.

The DNA sequence (SEQ ID NO:5) and derived amino acid

sequence (SEQ ID NO:6) of this specifically amplified DNA fragment of V1-1 are shown in the SEQUENCE Listings.

Nucleotides #1-#26 of SEQ ID NO:5 comprise a portion of oligonucleotide #1 and nucleotides #103 - #128 comprise a portion of the reverse compliment of oligonucleotide #2 utilized to perform the specific amplification reaction. Due to the function of oligonucleotides #1 and #2 in initiating the amplification reaction, they may not correspond exactly to the actual sequence encoding a V1-1 protein and are therefore not translated in the corresponding amino acid derivation (SEQ ID NO:6).

DNA sequence analysis of another subclone indicates that the specifically amplified DNA product contained therein encodes a portion of another BMP/TGF-β/Vg-1 (V1-1 related) protein of the invention named VL-1.

The DNA sequence (SEQ ID NO:7) and derived amino acid sequence (SEQ ID NO:8) of this specifically amplified DNA fragment are shown in the Sequence Listings.

Nucleotides #1 - #26 of SEQ ID NO:7 comprise a portion of oligonucleotide #1 and nucleotides #103 - #128 comprise a portion of the reverse compliment of oligonucleotide #2 utilized to perform the specific amplification reaction. Due to the function of oligonucleotides #1 and #2 in initiating the amplification reaction, they may not correspond exactly to the actual sequence encoding a VL-1 protein of the invention and are therefore not translated in the corresponding amino acid derivation (SEQ ID NO:8).

The following oligonucleotide probe is designed on the basis of the specifically amplified V1-1 human DNA sequence set forth above (SEQ ID NO:5) and synthesized on an automated DNA synthesizer:

#3: CCACTGCGAGGGCCTTTGCGACTTCCCTTTGCGTTCGCAC (SEQ ID NO:20)

This oligonucleotide probe is radioactively labeled with ^{32}P and employed to screen a human genomic library constructed

in the vector λFIX (Stratagene catalog #944201). 500,000 recombinants of the human genomic library are plated at a density of approximately 10,000 recombinants per plate on 50 plates. Duplicate nitrocellulose replicas of the recombinant bacteriophage plaques and hybridized to oligonucleotide probe #3 in standard hybridization buffer (SHB = 5X SSC, 0.1% SDS, 5X Denhardt's, 100 μg/ml salmon sperm DNA) at 65°C overnight. The following day the radioactively labelled oligonucleotide containing hybridization solution is removed an the filters are washed with 0.2X SSC, 0.1% SDS at 65°C. A single positively hybridizing recombinant is identified and plaque purified. This plaque purified recombinant bacteriophage clone which hybridizes to the V1-1 oligonucleotide probe #3 is designated λHuG-48. A bacteriophage plate stock is made and bacteriophage DNA is isolated from the λHuG-48 human genomic clone. The bacteriophage λHuG-48 has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under the accession #75625 on December 7, 1993. This deposit meets the requirements of the Budapest Treaty of the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations thereunder.

The oligonucleotide hybridizing region of this recombinant, λHuG-48, is localized to a 3.2 kb BamHI fragment. This fragment is subcloned into a plasmid vector (pGEM-3) and DNA sequence analysis is performed. This plasmid subclone is designated PCR1-1#2 and has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD "ATCC" under the accession #69517 on December 7, 1993. This deposit meets the requirements of the Budapest Treaty of the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and Regulations thereunder.

The partial DNA sequence (SEQ ID NO:1) and derived amino acid sequence (SEQ ID NO:2) of the 3.2 kb DNA insert of the plasmid subclone PCR1-1#2, derived from clone λHuG-48, are

shown in the Sequence Listings.

It should be noted that nucleotides #639 - #714 of SEQ ID NO:1 correspond to nucleotides #27 - #102 of the specifically amplified V1-1 encoding DNA fragment set forth in SEQ ID NO:5 thus confirming that the human genomic bacteriophage clone λHuG-48 and derivative subclone PCR1-1#2 encode at least a portion of the V1-1 protein of the invention. The nucleotide sequence of a portion of the 3.2 kb BamHI insert of the plasmid PCR1-1#2 contains an open reading frame of at least 882 base pairs, as defined by nucleotides #1-#882 of SEQ ID NO:1.

This open reading frame encodes at least 294 amino acids of the human V1-1 protein of the invention. The encoded 294 amino acid human V1-1 protein includes the full mature human V1-1 protein (amino acids #1-#104 of SEQ ID NO:2), as well as the C-terminal portion of the propeptide region of the primary translation product (amino acid #-190 to #-1 of SEQ ID NO:2).

Based on the knowledge of other BMP proteins and other proteins within the TGF-β family, it is predicted that the precursor polypeptide would be cleaved at the multibasic sequence Arg-Arg-Gly-Arg in agreement with a proposed consensus proteolytic processing sequence of Arg-X-X-Arg. Cleavage of the V1-1 precursor poly peptide is expected to generate a 104 amino acid mature peptide beginning with the amino acid Ser at position #1 of SEQ ID NO:2. The processing of V1-1 into the mature form is expected to involve dimerization and removal of the N-terminal region in a manner analogous to the processing of the related protein TGF-β [Gentry et al., Molec & Cell. Biol., 8:4162 (1988); Derynck et al. Nature, 316:701 (1985)].

It is contemplated therefore that the mature active species of V1-1 comprises a homodimer of two polypeptide subunits, each subunit comprising amino acids #1 to #104 of SEQ ID NO:2 with a predicted molecular weight of approximately 12,000 daltons. Further active species are contemplated comprising at least amino acids #3 to #103 of SEQ ID NO:2, thereby including the first conserved cysteine residue. As

with other members of the TGF-β/BMP family of proteins, the carboxy-terminal portion of the V1-1 protein exhibits greater sequence conservation than the more amino-terminal portion. The percent amino acid identity of the human V1-1 protein in the cysteine-rich C-terminal domain (amino acids #3 - #104) to the corresponding region of human BMP proteins and other proteins within the TGF-β family is as follows: BMP-2, 55%; BMP-3, 43%; BMP-4, 53%; BMP-5, 49%; BMP-6, 49%; BMP-7, 50%; BMP-8, 57%; BMP-9, 48%; BMP-10, 57%; activin WC, 38%; Vg1, 46%; GDF-1, 47%; TGF-β1, 36%; TGF-β2, 36%; TGF-β3, 39%; inhibin β(B), 36%; inhibin β(A), 41%.

The human V1-1 DNA sequence (SEQ ID NO:1), or a portion thereof, can be used as a probe to identify a human cell line or tissue which synthesizes V1-1 mRNA. Briefly described, RNA is extracted from a selected cell or tissue source and either electrophoresed on a formaldehyde agarose gel and transferred to nitrocellulose, or reacted with formaldehyde and spotted on nitrocellulose directly. The nitrocellulose is then hybridized to a probe derived from the coding sequence of human V1-1.

Alternatively, the human V1-1 sequence is used to design oligonucleotide primers which will specifically amplify a portion of the V1-1 encoding sequence located in the region between the primers utilized to perform the specific amplification reaction. It is contemplated that these human V1-1 derived primers would allow one to specifically amplify corresponding V1-1 encoding sequences from mRNA, cDNA or genomic DNA templates. Once a positive source has been identified by one of the above described methods, mRNA is selected by oligo (dT) cellulose chromatography and cDNA is synthesized and cloned in λgt10 or other λ bacteriophage vectors known to those skilled in the art, for example, λZAP by established techniques (Toole et al., supra). It is also possible to perform the oligonucleotide primer directed amplification reaction, described above, directly on a pre-established human cDNA or genomic library which has been cloned

into a λ bacteriophage vector. In such cases, a library which yields a specifically amplified DNA product encoding a portion of the human V1-1 protein could be screened directly, utilizing the fragment of amplified V1-1 encoding DNA as a probe.

Oligonucleotide primers designed on the basis of the DNA sequence of the human V1-1 genomic clone λHuG-48 are predicted to allow the specific amplification of human V1-1 encoding DNA sequences from pre-established human cDNA libraries which are commercially available (ie. Stratagene, La Jolla, CA or Clontech Laboratories, Inc., Palo Alto, CA). The following oligonucleotide primer is designed on the basis of nucleotides #571 - #590 of the DNA sequence set forth in SEQ ID NO:1 and synthesized on an automated DNA synthesizer:

#4: TGCGGATCCAGCCGCTGCAGCCGCAAGCC (SEQ ID NO:21)

The first nine nucleotides of primer #4 (underlined) comprise the recognition sequence for the restriction endonuclease BamHI which can be used to facilitate the manipulation of a specifically amplified DNA sequence encoding the human V1-1 protein of the invention and are thus not derived from the DNA sequence presented in SEQ ID NO:1.

The following oligonucleotide primer is designed on the basis of nucleotides #866 - #885 of the DNA sequence set forth in SEQ ID NO:1 and synthesized on an automated DNA synthesizer:

#5 GACTCTAGACTACCTGCAGCCGCAGGCCT (SEQ ID NO:22)

The first nine nucleotides of primer #5 (underlined) comprise the recognition sequence for the restriction endonuclease XbaI which can be used to facilitate the manipulation of a specifically amplified DNA sequence encoding the human V1-1 protein of the invention and are thus not derived from the DNA sequence presented in SEQ ID NO:1.

The standard nucleotide symbols in the above identified primers are as follows: A, adenine; C, cytosine; G, guanine; T, thymine.

Primers #4 and #5 identified above are utilized as primers to allow the amplification of a specific V1-1 encoding

nucleotide sequence from pre-established cDNA libraries which may include the following: human fetal brain cDNA/λZAPII (Stratagene catalog #936206), human liver/λUNI-ZAP XR (Stratagene Catalog #937200), human lung/λUNI-ZAP XR (Stratagene catalog #937206), and human fetal spleen/UNI-ZAP XR (Stratagene catalog #937205).

Approximately 1 x 10^8 pfu (plaque forming units) of λbacteriophage libraries containing human cDNA inserts such as those detailed above are denatured at 95°C for five minutes prior to addition to a reaction mixture containing 200 μM each deoxynucleotide triphosphates (dATP, dGTP, dCTP and dTTP) 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 1.25 units Taq DNA polymerase, 100 pM oligonucleotide primer #4 and 100 pM oligonucleotide primer #5. The reaction mixture is then subjected to thermal cycling in the following manner: 1 minute at 94°C, 1 minute at 50°C, 1 minute at 72°C for thirty-nine cycles followed by 10 minutes at 72°C.

The DNA which is specifically amplified by this reaction would be expected to generate a V1-1 encoding product of approximately 333 base pairs, the internal 315 bp of which correspond to nucleotides #571 - #885 of SEQ ID NO:1 and also including 9 bp at each end of the V1-1 specific fragment which correspond to the restriction sites defined by nucleotides #1 - #9 of primers #4 and #5. The resulting 333 bp DNA product is digested with the restriction endonucleases BamHI and XbaI, phenol extracted, chloroform extracted and ethanol precipitated.

Alternatively, to ethanol precipitation, buffer exchange and removal of small fragments of DNA resulting from the BamHI/XbaI restriction digest is accomplished by dilution of the digested DNA product in 10 mM Tris-HCl pH 8.0, 1 mM EDTA followed by centrifugation through a Centricon™ 30 microconcentrator (W.R. Grace & Co., Beverly, MA; Product #4209). The resulting BamHI/XbaI digested amplified DNA

product is subcloned into a plasmid vector (ie. pBluescript, pGEM-3 etc.) between the BamHI and XbaI sites of the polylinker region. DNA sequence analysis of the resulting subclones would be required to confirm the integrity of the V1-1 encoding insert. Once a positive cDNA source has been identified in this manner, the corresponding cDNA library from which a 333 bp V1-1 specific sequence was amplified could be screened directly with the 333 bp insert or other V1-1 specific probes in order to identify and isolate cDNA clones encoding the full-length V1-1 protein of the invention.

Additional methods known to those skilled in the art may be used to isolate other full-length cDNAs encoding human V1-1 related proteins, or full length cDNA clones encoding V1-1 related proteins of the invention from species other than humans, particularly other mammalian species.

The following examples demonstrate the use of the human V1-1 sequence to isolate homologues from V1-1 related proteins in a murine genomic DNA library.

The DNA sequence which encodes the human V1-1 protein of the invention is predicted to be significantly homologous to V1-1 and V1-1 related sequences from species other than humans that it could be utilized to specifically amplify DNA sequences from those other species which would encode the corresponding V1-1 related proteins. Specifically, the following oligonucleotides are designed on the basis of the human V1-1 sequence (SEQ ID NO:1) and are synthesized on an automated DNA synthesizer:

#6: GCGGATCCAAGGAGCTCGGCTGGGACGA (SEQ ID NO:23)

#7: GGAATTCCCCACCACCATGTCCTCGTAT (SEQ ID NO:24)

The first eight nucleotides of oligonucleotide primers #6 and #7 (underlined) comprise the recognition sequence for the restriction endonucleases BamHI and EcoRI, respectively. These sequences are utilized to facilitate the manipulation of a specifically amplified DNA sequence encoding a V1-1 or V1-1 related protein from a species other than human and are thus

not derived from the DNA sequence presented in SEQ ID NO:1.

Oligonucleotide primer #6 is designed on the basis of nucleotides #607-#626 of SEQ ID NO:1. Oligonucleotide primer #7 is designed on the basis of the reverse compliment of nucleotides #846-#865 of the DNA sequence set forth in SEQ ID NO:1.

Oligonucleotide primers #6 and #7 identified above are utilized as primers to allow the amplification of specific V1-1 related sequences from genomic DNA derived from species other than humans. The amplification reaction is performed as follows:

Murine genomic DNA (source: strain Balb c) is sheared by repeated passage through a 25 gauge needle, denatured at 100° C for five minutes and then chilled on ice before adding to a reaction mixture containing 200 μM each deoxynucleotide triphosphates (dATP, DGTP, dCTP and dTTP) 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 1.25 units Taq DNA polymerase, 100 pM oligonucleotide primer #6 and 100 pM oligonucleotide primer #7. The reaction mixture is then subjected to thermal cycling in the following manner: 1 minute at 95°C, 1 minute at 55°C, 1 minute at 72°C for forty cycles followed by 10 minutes at 72°C.

The DNA which is specifically amplified by this reaction is ethanol precipitated, digested with the restriction endonucleases BamHI and EcoRI and subjected to agarose gel electrophoresis. A region of the gel, corresponding to the predicted size of the murine V1-1 or V1-1 related encoding DNA fragment, is excised and the specifically amplified DNA fragments contained therein are extracted (by electroelution or by other methods known to those skilled in the art) and subcloned in to a plasmid vector, such as pGEM-3 or pBluescript between the BamHI and EcoRI sites of the polylinker. DNA sequence analysis of one of the resulting subclones named mV1, indicates that the specifically amplified DNA sequence

contained therein encodes a portion of a protein which appears to be the murine homolog to either the V1-1 or VL-1 sequence of the invention. The DNA sequence (SEQ ID NO:10) and derived amino acid sequence (SEQ ID NO:11) of this specifically amplified murine DNA fragment are shown in the sequence listings.

Nucleotides #1-#26 of SEQ ID NO:10 comprise a portion of oligonucleotide #6 and nucleotides #246-#272 comprise a portion of the reverse compliment of oligonucleotide #7 utilized to perform the specific amplification reaction. Nucleotide #27 of SEQ ID NO:10 appears to be the last nucleotide of a codon triplet, and nucleotides #244-#245 of SEQ ID NO:10 appear to be the first two nucleotides of a codon triplet. Therefore, nucleotides #28 to #243 of SEQ ID NO:10 correspond to a partial coding sequence of mV1. Due to the function of oligonucleotides #6 and #7 in initiating the amplification reaction, they may not correspond exactly to the actual sequence encoding the murine homolog to the human V1-1 or VL-1 protein of the invention and are therefore not translated in the corresponding amino acid sequence derivation (SEQ ID NO:11).

Oligonucleotide probes designed on the basis of the specifically amplified murine V1-1 or VL-1 DNA sequence set forth in SEQ ID NO:10 can be utilized by those skilled in the art to identify full-length murine V1-1 or VL-1 encoding clones (either cDNA or genomic).

DNA sequence analysis of another of the resulting subclones named mV2, indicates that the specifically amplified DNA sequence contained therein encodes a portion of a murine V1-1 related sequence of the invention. The DNA sequence (SEQ ID NO:12) and derived amino acid sequence (SEQ ID NO:13) of this specifically amplified murine DNA fragment are shown in the sequence listings.

Nucleotides #1-#26 of SEQ ID NO:12 comprise a portion of oligonucleotide #6 and nucleotides #246-#272 comprise a portion

of the reverse compliment of oligonucleotide #7 utilized to perform the specific amplification reaction. Nucleotide #27 of SEQ ID NO:12 appears to be the last nucleotide of a codon triplet, and nucleotides #244-#245 of SEQ ID NO:12 appear to be the first two nucleotides of a codon triplet. Therefore, nucleotides #28 to #243 of SEQ ID NO:12 correspond to a partial coding sequence of mV2. Due to the function of oligonucleotides #6 and #7 in initiating the amplification reaction, they may not correspond exactly to the actual sequence encoding the murine V1-1 related protein of the invention and are therefore not translated in the corresponding amino acid sequence derivation (SEQ ID NO:13).

Oligonucleotide probes designed on the basis of the specifically amplified murine V1-1 related DNA sequence set forth in SEQ ID NO:12 can be utilized by those skilled in the art to identify full-length murine V1-1 related encoding clones (either cDNA or genomic).

DNA sequence analysis of another of the resulting subclones named mV9, indicates that the specifically amplified DNA sequence contained therein encodes a portion of a murine V1-1 related sequence of the invention. This sequence appears to be the murine homolog to the human MP52 DNA sequence described at SEQ ID NO:3. The DNA sequence (SEQ ID NO:14) and derived amino acid sequence (SEQ ID NO:15) of this specifically amplified murine DNA fragment are shown in the sequence listings.

Nucleotides #1-#26 of SEQ ID NO:14 comprise a portion of oligonucleotide #6 and nucleotides #246-#272 comprise a portion of the reverse compliment of oligonucleotide #7 utilized to perform the specific amplification reaction. Nucleotide #27 of SEQ ID NO:14 appears to be the last nucleotide of a codon triplet, and nucleotides #244-#245 of SEQ ID NO:14 appear to be the first two nucleotides of a codon triplet. Therefore, nucleotides #28 to #243 of SEQ ID NO:14 correspond to a partial coding sequence of mV9. Due to the function of

oligonucleotides #6 and #7 in initiating the amplification reaction, they may not correspond exactly to the actual sequence encoding the murine V1-1 related protein of the invention and are therefore not translated in the corresponding amino acid sequence derivation (SEQ ID NO:15).

Oligonucleotide probes designed on the basis of the specifically amplified murine V1-1 related DNA sequence set forth in SEQ ID NO:14 can be utilized by those skilled in the art to identify full-length murine V1-1 related encoding clones (either cDNA or genomic).

In a similar manner to that which is described above for identifying and isolating human genomic clones encoding the V1-1 protein of the invention, oligonucleotide probe(s) corresponding to the VL-1 encoding sequence set forth in SEQ ID NO:7 can be designed and utilized to identify human genomic or cDNA sequences encoding the VL-1 protein. These oligonucleotides would be designed to regions specific for VL-1 encoding sequences and would therefore be likely to be derived from regions of the lowest degree of nucleotide sequence identity between the specifically amplified VL-1 encoding sequence (SEQ ID NO:7) and the specifically amplified V1-1 encoding sequence (SEQ ID NO:5).

EXAMPLE 2

Expression of V1-1

In order to produce human V1-1 proteins, the DNA encoding it is transferred into an appropriate expression vector and introduced into mammalian cells or other preferred eukaryotic or prokaryotic hosts by conventional genetic engineering techniques.

In order to produce the human V1-1 protein in bacterial cells, the following procedure is employed.

Expression of V1-1 in *E. coli*

An expression plasmid pALV1-781, for production of V1-1 in *E. coli* was constructed which contains the following principal features. Nucleotides 1-2060 contain DNA sequences originating

from the plasmid pUC-18 [Norrander et al., Gene 26:101-106 (1983)] including sequences containing the gene for β-lactamase which confers resistance to the antibiotic ampicillin in host *E. coli* strains, and a colE1-derived origin of replication. Nucleotides 2061-2221 contain DNA sequences for the major leftward promotor (pL) of bacteriophage λ [Sanger et al., J. Mol. Biol. 162:729-773 (1982)], including three operator sequences O_L1, O_L2 and O_L3. The operators are the binding sites for λcI repressor protein, intracellular levels of which control the amount of transcription initiation from pL. Nucleotides 2222-2723 contain a strong ribosome binding sequence included on a sequence derived from nucleotides 35566 to 35472 and 38137 to 38361 from bacteriophage lambda as described in Sanger et al., J. Mol. Biol. 162:729-773 (1982). Nucleotides 2724-3041 contain a DNA sequence encoding mature V1-1 protein with all 3' untranslated sequence removed. The V1-1 DNA sequences introduced into the pALV1-781 expression vector were modified at the 5'end to raise the A+T content without altering the coding capacity. These changes were made to increase the efficiency of translation initiated on the V1-1 mRNA in *E. coli*. Nucleotides 3042-3058 provide a "Linker" DNA sequence containing restriction endonuclease sites. Nucleotides 3059-3127 provide a transcription termination sequence based on that of the *E. coli* asp A gene [Takagi et al., Nucl. Acids Res. 13:2063-2074 (1985)]. Nucleotides 3128-3532 are DNA sequences derived from pUC-18.

Plasmid pALV1-781 was transformed into the *E. coli* host strain GI724 (F, lacI^q, lacp^{L8}, ampC::λcI^+) by the procedure of Dagert and Ehrlich, Gene 6:23 (1979). GI724 (ATCC accession No. 55151) contains a copy of the wild-type λcI repressor gene stably integrated into the chromosome at the ampC locus, where it has been placed under the transcriptional control of *Salmonella typhimurium* trp promotor/operator sequences. In GI724, λCI protein is made only during growth in tryptophan-

free media, such as minimal media or a minimal medium supplemented with casamino acids such as IMC, described above. Addition of tryptophan to a culture of GI724 will repress the trp promoter and turn off synthesis of λcI, gradually causing the induction of transcription from pL promoters if they are present in the cell.

Transformants were selected on 1.5% w/v agar plates containing IMC medium, which is composed of M9 medium [Miller, "Experiments in Molecular Genetics," Cold Spring Harbor Laboratory, New York (1972)] containing 1 mM $MgSO_4$ and supplemented with 0.5% w/v glucose, 0.2% w/v casamino acids and 100 μg/ml ampicillin. GI724 transformed with pALV1-781 was grown at 37°C to an A_{550} of 0.5 in IMC medium containing 100 μg/ml ampicillin. Tryptophan was then added to a final concentration of 100 μg/ml and the culture incubated for a further 4 hours. During this time V1-1 protein accumulates within the "inclusion body" fraction.

Preparation of Protein Monomer

18 g of frozen cells were weighed out and resuspended in 60ml of 100 mM Tris, 10 mM EDTA, 1 mM phenylmethylsulfonyl fluoride [PMSF], pH 8.3. Cells were lysed by 3 passes through a Microfluidizer™ [model #MCF 100 T]. The inclusion body pellet was obtained by centrifugation at 15,000g at 4°C for 20 minutes. The supernatant was decanted, and the pellet was washed with 100 ml of 100 mM Tris, 1.0 M NaCl, 10 mM EDTA, 1 mM PMSF, pH 8.3. The suspension was centrifuged again at 15,000g at 4°C for 10 minutes, and the supernatant decanted. The pellet was then washed with 100 ml of 100 mM Tris, 10 mM EDTA, 1% Triton X-100, 1 mM PMSF, pH 8.3. The suspension was centrifuged again at 15,000g at 4°C for 10 minutes, and the supernatant decanted. The pellet was resuspended with 50 ml of 20 mM Tris, 1 mM EDTA, 1 mM PMSF, pH 8.3, containing 1% DTT in a glass tissue homogenizer. Monomeric V1-1 was then solubilized by acidification to pH 2.5 with glacial acetic

acid. The soluble fraction was isolated by centrifugation at 15,000g for 20 minutes at 4°C.

The supernatant from this centrifugation was collected and chromatographed over a Sephacryl S-100™ size exclusion column (83 cm x 2.6 cm; ≈440 ml bed) in 20 ml increments. The Sephacryl S-100™ column was run with a mobile phase of 1% acetic acid at a flow rate of 1.4 ml/min. Fractions corresponding to V1-1 monomer were detected by absorbance at 280 nm, and using a computer calculated extinction coefficient of $18200 M^{-1}cm^{-1}$ and molecular weight (11667 daltons). This size exclusion column pooled material was used as starting material for refolding reactions.

As an alternative to the above, 1.0 g of cells stored at -80°C are measured. Solution (3.4 ml 100 mM TRIS, 10 mM EDTA, pH 8.5) is added. The solution is vortexed until cells are well suspended. 40 μl 100 mM PMSF in isopropanol is added. The cells are lysed at 1000 psi in a French pressure cell. The inclusion bodies are centrifuged at 4°C for 20 minutes in an Eppendorf microfuge to form pellets. The supernatants are decanted. To one pellet (out of 4 total) 1.0 ml degassed 8.0 M guanidine hydrochloride, 0.5 M TRIS, 5 mM EDTA, pH 8.5, containing 250 mM DTT is added. The pellet is dissolved and argon is blown over the liquid for 30 seconds. Next the solution is incubated at 37°C for one hour. Insoluble material is pelleted for 2-3 minutes in an Eppendorf microfuge at 23°C. 0.5-1.0 ml of supernatant is injected onto a Supelco 2 cm guard cartridge (LC-304), and eluted with an acetonitrile gradient in 0.1% TFA from 1-70% over 35 minutes. V1-1 elutes between 29 and 31 minutes. Fractions are pooled and the protein concentration determined by adsorbance at 280 nanometers versus 0.1% TFA, using the theoretical extinction coefficient based upon the amino acid content.

As a second alternate method to the above, frozen cell pellets obtained from the *E. coli* transformants as described

above are thawed in 30 ml of TE8.3(100:10) buffer (100 mM Tris-HCl pH 8.3, 10 mM Na_2EDTA, 1 mM PMSF). Cells are lysed by three passes through a Microfluidizer™ [model #MCF 100 T]. The initial inclusion body material pellet is dissolved in 8 M guanidine-HCl, TE8.5(100:10) buffer (100 mM Tris-HCl pH 8.5, 10 mM Na_2EDTA which contained 100 mM DTT, and incubated at 37°C for 1 hour. This material is centrifuged at 12,000 x g for 15 minutes at room temperature.

Refolding of V1-1 protein using CHAPS system

A sufficient volume of the V1-1 pool is lyophilized to give 10 μg of protein. 5 μl of glass distilled water is added to redissolve the residue, then 100 μl of refold mix (50 mM Tris, 1.0 M NaCl, 2% 3-(3-chlolamido-propyl)dimethylammonio-1-propane-sulfate (CHAPS), 5 mM EDTA, 2 mM glutathione (reduced) 1 mM glutathione (oxidized); at pH of approximately 8.5). The solution is gently mixed and stored at 23°C for 1-4 days. Dimer formation is assessed by running an aliquot on a Novex 16% tricine gel at 125 volts for 2.5 hours, followed by Coomassie Blue staining and destaining.

V1-1 dimer was purified using a C4 analytical RP-HPLC (reversed phase-high performance liquid chromatography) column (Vydac 214TP54) which was equilibrated to 1% B buffer (diluted into A buffer) and was run over 35 minutes, during which the protein elutes, using the following gradient (A buffer = 0.1% trifluoroacetic acid, B buffer = 95% acetonitrile, 0.1% trifluoroacetic acid [TFA]), with a flow rate of 1 ml/min:

1-5 minutes	20% B buffer
5-10 minutes	20-30% B buffer
10-30 minutes	30-50% B buffer
30-35 minutes	50-100% B buffer

Protein was monitored by absorbance at 280nm. Peak V1-1 fractions (eluting between 29 and 31 minutes) were pooled. Purity was assessed by SDS-PAGE. The concentration was determined by absorbance at 280nm, and using the computer

calculated extinction coefficient and molecular weight as indicated above.

Expression of V1-1 in mammalian cells:

Another contemplated preferred expression system for biologically active recombinant human V1-1 is stably transformed mammalian cells.

One skilled in the art can construct mammalian expression vectors by employing the sequence of SEQ ID NO:1, or other DNA sequences encoding V1-1 proteins or other modified sequences and known vectors, such as pCD [Okayama et al., Mol. Cell Biol., 2:161-170 (1982)], pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)] and pMT2 CXM.

The mammalian expression vector pMT2 CXM is a derivative of p91023(b) (Wong et al., Science 228:810-815, 1985) differing from the latter in that it contains the ampicillin resistance gene in place of the tetracycline resistance gene and further contains a XhoI site for insertion of cDNA clones. The functional elements of pMT2 CXM have been described (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82:689-693) and include the adenovirus VA genes, the SV40 origin of replication including the 72 bp enhancer, the adenovirus major late promoter including a 5′ splice site and the majority of the adenovirus tripartite leader sequence present on adenovirus late mRNAs, a 3′ splice acceptor site, a DHFR insert, the SV40 early polyadenylation site (SV40), and pBR322 sequences needed for propagation in E. coli.

Plasmid pMT2 CXM is obtained by EcoRI digestion of pMT2-VWF, which has been deposited with the American Type Culture Collection (ATCC), Rockville, MD (USA) under accession number ATCC 67122. EcoRI digestion excises the cDNA insert present in pMT2-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods. pMT2 CXM is then constructed using loopout/in mutagenesis [Morinaga, et al., Biotechnology 84: 636 (1984).

This removes bases 1075 to 1145 relative to the Hind III site near the SV40 origin of replication and enhancer sequences of pMT2. In addition it inserts a sequence containing the recognition site for the restriction endonuclease Xho I. A derivative of pMT2CXM, termed pMT23, contains recognition sites for the restriction endonucleases PstI, Eco RI, SalI and XhoI. Plasmid pMT2 CXM and pMT23 DNA may be prepared by conventional methods.

pEMC2β1 derived from pMT21 may also be suitable in practice of the invention. pMT21 is derived from pMT2 which is derived from pMT2-VWF. As described above EcoRI digestion excises the cDNA insert present in pMT-VWF, yielding pMT2 in linear form which can be ligated and used to transform E. Coli HB 101 or DH-5 to ampicillin resistance. Plasmid pMT2 DNA can be prepared by conventional methods.

pMT21 is derived from pMT2 through the following two modifications. First, 76 bp of the 5′ untranslated region of the DHFR cDNA including a stretch of 19 G residues from G/C tailing for cDNA cloning is deleted. In this process, a XhoI site is inserted to obtain the following sequence immediately upstream from DHFR. Second, a unique ClaI site is introduced by digestion with EcoRV and XbaI, treatment with Klenow fragment of DNA polymerase I, and ligation to a ClaI linker (CATCGATG). This deletes a 250 bp segment from the adenovirus associated RNA (VAI) region but does not interfere with VAI RNA gene expression or function. pMT21 is digested with EcoRI and XhoI, and used to derive the vector pEMC2B1.

A portion of the EMCV leader is obtained from pMT2-ECAT1 [S.K. Jung, et al, J. Virol 63:1651-1660 (1989)] by digestion with Eco RI and PstI, resulting in a 2752 bp fragment. This fragment is digested with TaqI yielding an Eco RI-TaqI fragment of 508 bp which is purified by electrophoresis on low melting agarose gel. A 68 bp adapter and its complementary strand are synthesized with a 5′ TaqI protruding end and a 3′ XhoI protruding end which has a sequence which matches the EMC virus

leader sequence from nucleotide 763 to 827. It also changes the ATG at position 10 within the EMC virus leader to an ATT and is followed by a XhoI site. A three way ligation of the pMT21 Eco RI-XhoI fragment, the EMC virus EcoRI-TaqI fragment, and the 68 bp oligonucleotide adapter TaqI-XhoI adapter resulting in the vector pEMC2β1.

This vector contains the SV40 origin of replication and enhancer, the adenovirus major late promoter, a cDNA copy of the majority of the adenovirus tripartite leader sequence, a small hybrid intervening sequence, an SV40 polyadenylation signal and the adenovirus VA I gene, DHFR and β-lactamase markers and an EMC sequence, in appropriate relationships to direct the high level expression of the desired cDNA in mammalian cells.

The construction of vectors may involve modification of the V1-1 DNA sequences. For instance, V1-1 cDNA can be modified by removing the non-coding nucleotides on the 5′ and 3′ ends of the coding region. The deleted non-coding nucleotides may or may not be replaced by other sequences known to be beneficial for expression. These vectors are transformed into appropriate host cells for expression of V1-1 proteins. Additionally, the sequence of SEQ ID NO:1 or other sequences encoding V1-1 proteins can be manipulated to express V1-1 protein by isolating the mature coding sequence of nucleotides 571 to 882 of SEQ ID NO:1 and adding at the 5′ end sequences encoding the complete propeptides of other BMP proteins.

One skilled in the art can manipulate the sequences of SEQ ID NO:1 by eliminating or replacing the mammalian regulatory sequences flanking the coding sequence with bacterial sequences to create bacterial vectors for intracellular or extracellular expression by bacterial cells, as described above. As another example, the coding sequences could be further manipulated (e.g. ligated to other known linkers or modified by deleting non-coding sequences therefrom or altering nucleotides therein by other known techniques). The modified V1-1 coding sequence

could then be inserted into a known bacterial vector using procedures such as described in T. Taniguchi et al., Proc. Natl Acad. Sci. USA, 77:5230-5233 (1980). This exemplary bacterial vector could then be transformed into bacterial host cells and a V1-1 protein expressed thereby. For a strategy for producing extracellular expression of V1-1 proteins in bacterial cells, see, e.g. European patent application EPA 177,343.

Similar manipulations can be performed for the construction of an insect vector [See, e.g. procedures described in published European patent application 155,476] for expression in insect cells. A yeast vector could also be constructed employing yeast regulatory sequences for intracellular or extracellular expression of the factors of the present invention by yeast cells. [See, e.g., procedures described in published PCT application WO86/00639 and European patent application EPA 123,289].

A method for producing high levels of a V1-1 protein of the invention in mammalian cells may involve the construction of cells containing multiple copies of the heterologous V1-1 gene. The heterologous gene is linked to an amplifiable marker, e.g. the dihydrofolate reductase (DHFR) gene for which cells containing increased gene copies can be selected for propagation in increasing concentrations of methotrexate (MTX) according to the procedures of Kaufman and Sharp, J. Mol. Biol., 159:601-629 (1982). This approach can be employed with a number of different cell types.

For example, a plasmid containing a DNA sequence for a V1-1 of the invention in operative association with other plasmid sequences enabling expression thereof and the DHFR expression plasmid pAdA26SV(A)3 [Kaufman and Sharp, Mol. Cell. Biol., 2:1304 (1982)] can be co-introduced into DHFR-deficient CHO cells, DUKX-BII, by various methods including calcium phosphate coprecipitation and transfection, electroporation or protoplast fusion. DHFR expressing transformants are selected for growth in alpha media with dialyzed fetal calf serum, and subsequently

selected for amplification by growth in increasing concentrations of MTX (e.g. sequential steps in 0.02, 0.2, 1.0 and 5uM MTX) as described in Kaufman et al., Mol Cell Biol., 5:1750 (1983). Transformants are cloned, and biologically active V1-1 expression is monitored by the Rosen-modified Sampath-Reddi rat assay described below in Example 4. V1-1 expression should increase with increasing levels of MTX resistance. V1-1 polypeptides are characterized using standard techniques known in the art such as pulse labeling with [35S] methionine or cysteine and polyacrylamide gel electrophoresis. Similar procedures can be followed to produce other related V1-1 proteins.

EXAMPLE 3

Biological Activity of Expressed V1-1

To measure the biological activity of the expressed V1-1 proteins obtained in Example 2 above, the proteins are recovered from the cell culture and purified by isolating the V1-1 proteins from other proteinaceous materials with which they are co-produced as well as from other contaminants. The purified protein may be assayed in accordance with the rat assay described below in Example 4.

Purification is carried out using standard techniques known to those skilled in the art.

Protein analysis is conducted using standard techniques such as SDS-PAGE acrylamide [Laemmli, Nature 227:680 (1970)] stained with Coomassie Blue or silver [Oakley, et al. Anal. Biochem. 105:361 (1980)] and by immunoblot [Towbin, et al. Proc. Natl. Acad. Sci. USA 76:4350 (1979)]

EXAMPLE 4

ROSEN MODIFIED SAMPATH-REDDI ASSAY

A modified version of the rat ectopic implant assay described in Sampath and Reddi, Proc. Natl. Acad. Sci. USA, 80:6591-6595 (1983) is used to evaluate the activity of the V1-1 proteins. This modified assay is herein called the Rosen-modified Sampath-Reddi assay. The assay has been widely used

to evaluate the bone and cartilage-inducing activity of BMPs. The ethanol precipitation step of the Sampath-Reddi procedure is replaced by dialyzing (if the composition is a solution) or diafiltering (if the composition is a suspension) the fraction to be assayed against water. The solution or suspension is then equilibrated to 0.1% TFA. The resulting solution is added to 20 mg of rat matrix. A mock rat matrix sample not treated with the protein serves as a control. This material is frozen and lyophilized and the resulting powder enclosed in #5 gelatin capsules. The capsules are implanted subcutaneously in the abdominal thoracic area of 21-49 day old male Long Evans rats. The implants are removed after 10 days.

A section of each implant is fixed and processed for histological analysis. 1 μm glycolmethacrylate sections are stained with Von Kossa and acid fuschin to score the amount of induced tendon/ligament-like tissue formation present in each implant.

V1-1 was implanted in the rates in doses of 1, 5, 25 and 50 μg per implant for 10 days. BMP-2 at a dose of 5 μg was included as a positive control. For all doses of V1-1 tested, no bone or cartilage formation was observed in the implants after ten days. Instead, the implants were filled with tissue resembling embryonic tendon, which is easily recognized by the presence of dense bundles of fibroblasts oriented in the same plane and packed tightly together. [Tendon/ligament-like tissue is described, for example, in Ham and Cormack, Histology (JB Lippincott Co. (1979), pp. 367-369, the disclosure of which is hereby incorporated by reference]. These findings were reproduced in a second set of assays in which tendon/ligament-like tissues was present in all V1-1 containing implants. In contrast, the BMP-2 implants, as expected, showed cartilage and bone formation, but contained no tendon/ligament-like tissue.

The V1-1 proteins of this invention may be assessed for activity on this assay.

The foregoing descriptions detail presently preferred

embodiments of the present invention. Numerous modifications and variations in practice thereof are expected to occur to those skilled in the art upon consideration of these descriptions. Those modifications and variations are believed to be encompassed within the claims appended hereto.